Commun Nonlinear Sci Numer Simulat 16 (2011) 4698–4703

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

journal homepage: www.elsevier.com/locate/cnsns

Complex valued neural network with Möbius activation function

Necati Özdemir *, Beyza B. İskender, Nihal Yılmaz Özgür

Department of Mathematics, Faculty of Science and Arts, Balıkesir University, Cagis Campus, 10145 Balıkesir, Turkey

A R T I C L E   I N F O

Article history:
Available online 21 March 2011

Keywords:
Complex valued neural networks
Möbius transformation
Reflection in a circle
Lyapunov stability

A B S T R A C T

In this work, we propose a new type of activation function for a complex valued neural network (CVNN). This activation function is a special Möbius transformation classified as reflection. It is bounded outside of the unit disk and has partial continuous derivatives but not differentiable since it does not satisfy the Cauchy–Riemann equalities. However, the fixed points set of this function is a circle. Therefore, we employ this function to a specific complex valued Hopfield neural network (CVHNN) and increase the number of fixed points of the CVHNN. Using of this activation function leads us also to guarantee the existence of fixed points of the CVHNN. It is shown that the fixed points are all asymptotically stable states of the CVHNN which indicates that the information capacity is enlarged.

1. Introduction

A complex valued neural network (CVNN) is a neural network that processes information in the complex plane $\mathbb{C}$ [1]. It becomes very attractive field at the end of the 1980s and applicable to optoelectronics, imaging, remote sensing, quantum neural devices and systems, spatiotemporal analysis of physiological neural systems, and artificial neural information processing, see [2].

For CVNNs, the main task is to find a suitable activation function in a variety of complex functions. Despite the activation function of real valued neural networks (RVNNs) is chosen to be smooth and bounded generally as a sigmoid function, in the complex plane these properties are not convenient for the nature of neural networks. Because of Liouville's theorem; the analytic and bounded functions on entire complex plane are constant.

There are several complex activation functions proposed in the literature. The basic ones of them are given below.

The sigmoid function was also used for CVNNs by Leung and Haykin [3]

$$f(z) = \frac{1}{1+e^{-z}},$$

but this function has singular points at every $z = (2n+1)i\pi,\ n \in \mathbb{Z}$. They avoided this problem by scaling the input data to some region of the complex plane. Later on, the sigmoid function was adapted to CVNNs as

$$f(z) = \frac{1}{1+e^{-Rez}} + i\frac{1}{1+e^{-Imz}}$$

by Birx and Pipenberg [4]; Benvenuto and Piazza [5]. Also, *tanh* function which has singular points at every $z = (n+\frac{1}{2})i\pi,\ n \in \mathbb{Z}$ was adapted to CVNNs as real-imaginary type activation function

$$f(z) = \tanh(Rez) + i\tanh(Imz)$$

* Corresponding author. Tel.: +90 26661 121000x1215; fax: +90 2666 121215.
E-mail addresses: nozdemir@balikesir.edu.tr (N. Özdemir), biskender@balikesir.edu.tr (B.B. İskender), nihal@balikesir.edu.tr (N.Y. Özgür).

1007-5704/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cnsns.2011.03.005

by Kechriotis and Monalakos [6]; Kinouchi and Hagiwara [7], and as amplitude-phase type activation function

$$f(z) = \tanh(|z|)\exp(i\arg(z))$$

by Hirose [8].

The other activation functions are given below:

$$f(z) = \frac{z}{|z|}$$

by Noest [9],

$$f(z) = \frac{z}{c + \frac{1}{r}|z|}$$

by Georgiou and Koutsougeras [10],

$$f(z) = \frac{Rez}{c + \frac{1}{r}|Rez|} + i\frac{Imz}{c + \frac{1}{r}|Imz|}$$

or

$$f(z) = \frac{|z|}{c + \frac{1}{r}|z|}\exp\left[i\left\{\arg z - \frac{1}{2^n}\sin(2^n \arg z)\right\}\right]$$

by Kuroe and Taniguchi [11], in which c and $\frac{1}{r}$ are positive constants and $-\pi < \arg z < \pi$. Detailed comparison for these types of activation functions can be found in [12]. In addition, Kim and Adali [13] presented a set of elementary transcendental functions whose components are bounded almost everywhere and analytic functions to employ backpropagation. The *tanh* function is one of them and the singularities of the function was avoided by restricting the domain of interest to a circle of radius $\frac{\pi}{2}$.

Another approach to chose activation functions of CVNNs using conformal mappings was proposed by Clarke [14]. He emphasized that the elegant theory of conformal mappings can be applied to find other activation functions. He gave the following activation function

$$f(z) = \frac{(\cos\theta + i\sin\theta)(z - \alpha)}{1 - \bar{\alpha}z},$$

where θ is a rotation angle, α is a complex constant with $|\alpha| < 1$ and $\bar{\alpha}$ denotes complex conjugate of α. This function is the general conformal mapping that transform unit disk in the complex plane onto itself and also a Möbius transformation. Furthermore, Möbius transformations were used in RVNNs by Mandic [15]. He showed that sigmoidal or *tanh* types of activation functions for a RVNN satisfy the conditions of a Möbius transformation. To base on the observation of "fixed points of a neural network are determined by fixed points of the employed activation function" he deduced "the existence for fixed points of the activation function are guaranteed by the Möbius transformation".

In this work, we consider a new complex activation function known as reflection type Möbius transformation whose details are given in Section 2. Our motivation to chose this function is to enlarge information capacity of CVNNs. As it is known, information in a neural network is stored as asymptotically stable states [16]. The proposed function has infinite number of fixed points which lie on a circle and corresponds to fixed points of the considering CVNN in Section 3. We investigate stability of the fixed points in Section 4 by using Lyapunov stability approach and show that the fixed points are all asymptotically stable states of the CVNN under the assumptions of Theorem 2.

2. Möbius transformation as activation function

A Möbius transformation is defined as

$$f(z) = \frac{az + b}{cz + d}, \tag{1}$$

where $a,\ b,\ c,\ d \in \mathbb{C}$ and $ad - bc = 1$. It is a conformal mapping of the complex plane and also known as linear fractional or bilinear transformation. Such a Möbius transformation has at most two fixed points if it is not identity transformation $f(z) = z$. Detailed information could be found in [17,18].

Möbius transformations with real coefficients can be classified into

$$G_1 = \left\{f : f(z) = \frac{az + b}{cz + d},\ a,b,c,d \in \mathbb{R},\ ad - bc = 1\right\} \tag{2}$$

and

$$G_2 = \left\{g : g(z) = \frac{a\bar{z} + b}{c\bar{z} + d},\ a,b,c,d \in \mathbb{R},\ ad - bc = -1\right\}. \tag{3}$$

The transformations belong to the components of $G = G_1 \cup G_2$ are bijective transformations of extended complex plane. The transformations in G_1 are conformal mappings and have at most two fixed points. Any transformation belongs to G_2 is anti-conformal mapping and can be classified according to the value of $a + d$. If $a + d \neq 0$, the transformation is called as a glide-reflection and has two fixed points on the real axis. If $a + d = 0$, the transformation is called as a reflection and has infinite number of fixed points on a circle centered at $\frac{a}{c}$ and of radius $\frac{1}{|c|}$. To utilize the infinite number of fixed points, we use reflection type Möbius transformation as an activation function. We begin to analyse these types of activation functions by choosing a simple reflection transformation

$$f(z) = \frac{1}{\bar{z}}. \tag{4}$$

This transformation maps unit circle onto itself, outside of the unit circle to its inside and inside of the unit circle to its outside. It is not differentiable since it does not satisfy the Cauchy–Riemann equalities and has a singularity at $z = 0$. This transformation is bounded only for the points at the outside of the unit circle, see Fig. 1. Therefore, we restrict the domain of interest to the set of

$$B = \{z : |z| \geqslant 1\}. \tag{5}$$

Remark 1. Let γ be the circle centered at p and of radius r. Then it is known that the reflection transformation in the circle γ is denoted by $I_\gamma(z)$ and defined as follows:

$$I_\gamma(z) = \frac{r^2}{\bar{z} - \bar{p}} + p. \tag{6}$$

This indicates that any circle in the complex plane can be represented by a unique Möbius transformation, see [19]. When the circle is centered on real axis, the transformation is to be a reflection. Indeed, we have

$$I_\gamma(z) = \frac{p\bar{z} + r^2 - |p|^2}{\bar{z} - \bar{p}}.$$

Now, we divide the numerator and the denominator of this transformation with r, then we have $ad - bc = -1$ and $a + d = \frac{p}{r} - \frac{\bar{p}}{r} = 0$ since $p \in \mathbb{R}$. It means that we have the advantage of determining the fixed point circle in the complex plane. In this paper, we analyse the CVNN whose fixed points set is chosen as the circle centered at the origin with radius $r = 1$. The analysis is also valid for the circles that is centered at the origin and with any value of radius. From Eq. (6) the reflection transformation of a circle centered at the origin and of radius r is

$$f(z) = \frac{r^2}{\bar{z}}. \tag{7}$$

This transformation is bounded in the following set

$$B = \{z : |z| \geqslant r\}$$

and maps outside of the circle with radius r to inside of the circle with radius r. Thus, we can deduce that the domain of interest and the fixed points circle can be adjusted.

In the following section, we give a CVNN model to analyse the advantages of the new type activation function.

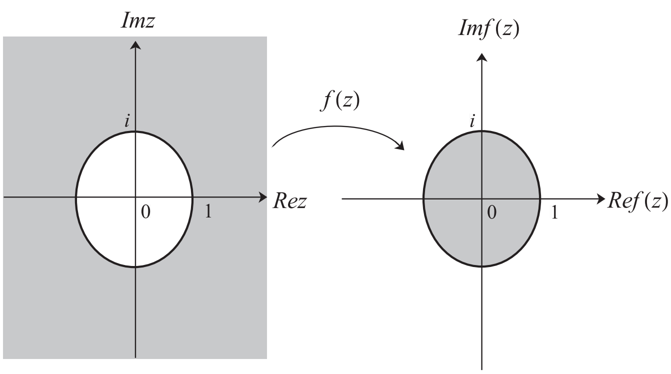

Fig. 1. Geometric interpretation of the reflection transformation given by Eq. (4).

3. Complex valued Hopfield neural network

Hopfield neural network can be considered as a class of nonlinear and autonomous system, see [20,21]. We consider this class of system in the complex plane in order to interest complex valued Hopfield neural network (CVHNN) given by

$$\dot{z}(t) = -H(z(t))(-Tz(t) + F(z(t)) - U),$$

where $T \in \mathbb{C}^{n\times n}$, $U \in \mathbb{C}^n$ are matrices, $z(t) \in \mathbb{C}^n$ is state vector, $H(z) : \mathbb{C}^n \to \mathbb{C}^{n\times n}$ is a nonlinear function and $F(z) = (f_1(z_1), f_2(z_2), \dots, f_n(z_n))^T : \mathbb{C}^n \to \mathbb{C}^n$ is an activation function. The activation function is chosen as in Eq. (4):

$$f_j(z_j) = \frac{1}{\bar{z}_j}, \quad j = 1, 2, \dots, n. \tag{8}$$

To obtain correspondence between fixed points of the activation function and fixed points of the network, we select $T \in \mathbb{R}^{n\times n}$ and $U = 0$. Hence, we interest the CVHNN with the form of

$$\dot{z}(t) = -H(z(t))(-Tz(t) + F(z(t))). \tag{9}$$

Fixed points of the Eq. (9) are calculated by the following equation:

$$-H(z)(-Tz + F(z)) = 0.$$

Assume that $H(z)$ is a nonsingular matrix then the fixed points are

$$F(z) = Tz$$

which correspond to the fixed points of the activation function.

4. Stability of fixed points

As mentioned in Section 3, information are stored in a neural network as asymptotically stable states. A stable state is a fixed point of the neural network and also known as equilibrium point. Therefore, it is important to increase the number of stable states in a neural network. By using the activation function in Eq. (8), we increase the number of fixed points, but now we must know whether the fixed points are stable or not. We investigate stability of the fixed points by using Lyapunov stability.

Definition 1. $E(z)$ is a Lyapunov function of the CVHNN if $E(z)$ is a mapping $E : \mathbb{C}^n \to \mathbb{R}$ and the derivative of E along the trajectory of CVHNN satisfies $\dot{E}(z) \leqslant 0$. Furthermore, $\dot{E}(z) = 0$ if and only if $\dot{z} = 0$.

If all equilibrium points of the network are isolated and CVHNN given by Eq. (9) has a Lyapunov function, then no nontrivial periodic solution exists and each solution of the network converges to an equilibrium point as $t \to \infty$, see [22].

An equilibrium point is isolated if it has no other equilibrium points in its vicinity, or there could be a continuum (compact and connected set) of equilibrium points, [23]. The fixed points of the CVHNN are isolated since they are on a circle. Therefore, the following theorem gives the stability of the fixed points. Here, we use the inner product defined on $\mathbb{C}^n$ as

$$\langle z_1, z_2 \rangle = z_2^* z_1,$$

where $z_1,\ z_2 \in \mathbb{C}^n$ and $(\cdot)^*$ denotes the conjugate transpose.

Theorem 2. *If the matrix $T \in \mathbb{R}^{n\times n}$ is symmetric and the matrix Re[H(z)] is positive definite, then the function*

$$E(z) = -\frac{1}{2} z^* T z + Re\left[\sum_{j=1}^{n} \int_0^{\bar{z}_j} \bar{f}_j(s) ds\right] \tag{10}$$

is a Lyapunov function of the CVHNN given by Eq. (9).

Proof. We can write Eq. (10) in the component wise form as

$$E(z) = -\frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \bar{z}_j T_{jk} z_k + \mathrm{Re}\left[\sum_{j=1}^{n} \int_0^{\bar{z}_j} \bar{f}_j(s) ds\right].$$

To show the monotonic decreasing of E with time t we compute $\dot{E}(z)$. Differentiating the first term of E gives

$$-\frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \left(\frac{d\bar{z}_j}{dt} T_{jk} z_k + \frac{dz_j}{dt} T_{jk} \bar{z}_k + \frac{d\bar{z}_k}{dt} T_{jk} z_j + \frac{dz_k}{dt} T_{jk} \bar{z}_j \right).$$

By using the symmetry property of the T matrix, this term can be arranged as follow

$$-\mathrm{Re}\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\left(T_{jk}z_k\frac{d\bar{z}_j}{dt}\right)\right].$$

Therefore,

$$\dot{E}(z) = -\mathrm{Re}\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\left(T_{jk}z_k\frac{d\bar{z}_j}{dt}\right)\right] + \mathrm{Re}\left[\sum_{j=1}^{n}\bar{f}_j(\bar{z}_j)\frac{d\bar{z}_j}{dt}\right].$$

Using the property of $\bar{f}_j(\bar{z}_j) = f_j(z_j)$, this equation can be written in the matrix form as

$$\dot{E}(z) = -\mathrm{Re}[(Tz - F(z))\dot{z}^*]. \tag{11}$$

Substituting $\dot{z}^*$ into the Eq. (11) gives

$$\begin{aligned}\dot{E}(z) &= -\mathrm{Re}[(Tz - F(z))(Tz - F(z))^* H(z)^*] \\ &= -\mathrm{Re}[(Tz - F(z))]^2 \mathrm{Re}[H(z)^*]\end{aligned}$$

which is negative for positive definite $\mathrm{Re}[H(z)]$ matrix and also equal to zero if and only if $\dot{z}(t) = 0$. □

Theorem 2 shows that the proposed activation function leads to infinite number of stable states. Consequently, number of the stored information is increased.

5. Conclusions

This paper is constructed on the idea of interesting geometric properties of Möbius transformations which are conformal mappings of the complex plane. Because of a special class of Möbius transformation defined in Eq. (3)

$$g(z) = \frac{a\bar{z} + b}{c\bar{z} + d}, \quad a, b, c, d \in \mathbb{R}, \quad ad - bc = -1$$

has infinite number of fixed points if $a + d = 0$, we think to combine this property with complex valued neural network (CVNN) and aim to increase the number of stored information in a CVNN. Thus, we have used a simple Möbius transformation in the type of reflection $f(z) = \frac{1}{\bar{z}}$ that maps the unit circle onto itself, its inside to its outside and vice versa. Since this function has singularity at the origin and unbounded in the unit circle, we have restricted the domain of interest of the CVNN to the outside of the unit circle. Therefore, we have guaranteed the boundedness of the function which is an important feature for neural networks. We have employed this activation function to a specific complex valued Hopfield neural network (CVHNN) and showed that the fixed points of the activation function are the fixed points of the CVHNN. Finally, we have proved that the fixed points are stable states for positive real valued function $\mathrm{Re}[H(z)] > 0$ which indicates that the information capacity is enlarged. In addition, it has been pointed out that the analysis is also valid for the activation functions in the form of

$$f(z) = \frac{r^2}{\bar{z}}, \quad r \in \mathbb{R}, \tag{12}$$

whose fixed points are on a circle with radius of r. This gives the opportunity of adjusting the domain of interest and the fixed points circle of the neural network.

References

[1] Hirose A. Complex valued neural networks. Berlin Heidelberg: Springer-Verlag; 2006.

[2] Gangal AS, Kalra PK, Chauhan DS. Performance evaluation of complex valued neural networks using various error functions. World Acad Sci Eng Technol 2007;29:27–32.

[3] Leung H, Haykin S. The complex backpropagation algorithm. IEEE Trans Signal Process 1991;39:2101–4.

[4] Birx DL, Pipenberg SJ. Chaotic oscillators and complex mapping feed forward networks (CMFFNS) for signal detection in noisy environments, Proceedings of IEEE international jt. conf. neural networks. II; 1992. p. 881–888.

[5] Benvenuto N, Piazza F. On the complex backpropagation algorithm. IEEE Trans Signal Process 1992;40:967–9.

[6] Kechriotis G, Monalakos ES. Training fully recurrent neural networks with complex weights. IEEE Trans Circuits Syst II 1994;41:235–8.

[7] Kinouchi M, Hagiwara M. Learning temporal sequences by complex neurons with local feedback, Proceedings of IEEE international conference on neural networks IV; 1995. p. 3165–3169.

[8] Hirose A. Applications of complex-valued neural networks to coherent optical computing using phase-sensitive detection scheme. Inf Sci Appl 1994;2:103–17.

[9] Noest AJ. Associative memory in sparse phasor neural networks. Europhys Lett 1988;6:469–74.

[10] Georgiou GM, Koutsougeras C. Complex domain backpropagation. IEEE Trans Circuits Syst II 1992;39:330–4.

[11] Kuroe Y, Taniguchi T. Models of self-correlation type complex-valued associative memories and their dynamics. In: Duch W et al., editors. Artificial neural networks: biological inspirations-ICANN 2005. Lecture notes in computer science, 3696. Springer-Verlag; 2005. p. 185–92.

[12] Kuroe Y, Taniguchi T. Models of complex-valued Hopfield-type neural networks and their dynamics. In: Nitta T, editor. Complex-valued neural networks, Information science reference, Hersey New York; 2009. p. 123–141.

[13] Kim T, Adali T. Fully complex multi-layer perceptron network for nonlinear signal processing. J VLSI Signal Process 2002;32:29–43.

[14] T. Clarke, Generalization of Neural Network to the Complex Plane, Proceedings of international jt. conference on neural networks, vol. 2; 1990. p. 435–440.
[15] Mandic DP. The use of Möbius transformations in neural networks and signal processing. Neural Netw Signal Process X 2000;1:185–94.
[16] Abu-Mostafa YS, Jacques J St. Information capacity of the Hopfield model. IEEE Trans Inf Theory IT-31 1985:461–4.
[17] Jones GA, Singerman D. Complex functions an algebraic and geometric viewpoint. Great Britain: Cambridge University Press; 1994.
[18] Needham T. Visual complex analysis. Great Britain: Oxford University Press; 2000.
[19] Brickman L. The symmetry principle for Möbius transformations. Am Math Mon 1993;100:781–2.
[20] Li JH, Michel AN, Porod W. Qualitative analysis and synthesis of a class of neural networks. IEEE Trans Circuits Syst 1988;35:976–86.
[21] Leblebicioğlu K, Halıcı U, Çelebi O. Infinite dimensional Hopfield neural networks. Nonlinear Anal Theory Methods Appl 2001;47:5807–13.
[22] Kuroe Y, Yoshida M, Mori T. On activation functions for complex-valued neural networks – existence of energy functions. In: Kaynak O et al., editors. ICANN/ICONIP 2003. LNCS 2714. Berlin Heidelberg: Springer-Verlag; 2003. p. 985–92.
[23] Khalil HK. Nonlinear systems. 2nd ed. United States of America: Prentice Hall; 1996.